BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Crystal structure of an HIV-binding recombinant fragment of human CD4

Seong-Eon Ryu*, Peter D. Kwong*, Alemseged Truneh†, Terence G. Porter†, James Arthos†, Martin Rosenberg†, Xiaoping Dai‡, Nguyen-huu Xuong‡, Richard Axel*§, Raymond W. Sweet† & Wayne A. Hendrickson*§

* Department of Biochemistry and Molecular Biophysics and § Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA
† SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406, USA
‡ Departments of Physics, Chemistry and Biology, University of California at San Diego, La Jolla, California 90293, USA

CD4 glycoprotein on the surface of T cells helps in the immune response and is the receptor for HIV infection. The structure of a soluble fragment of CD4 determined at 2.3 Å resolution reveals that the molecule has two intimately associated immunoglobulin-like domains. Residues implicated in HIV recognition by analysis of mutants and antibody binding are salient features in domain D1. Domain D2 is distinguished by a variation on the β-strand topologies of antibody domains and by an intra-sheet disulphide bridge.

CD4, a cell-surface glycoprotein found primarily on T lymphocytes, is required to shape the T-cell repertoire during thymic development and to permit appropriate activation of mature T cells[1]. T cells that recognize antigens associated with class II major histocompatibility complex (MHC) molecules, mainly T helper cells, express CD4. Evidence is accumulating that CD4 and the T-cell receptor coordinately engage class II molecules on antigen-presenting cells to mediate an efficient cellular immune response, and that engaged CD4 may transmit a signal to an associated cytoplasmic tyrosine kinase, p56^{lck}.

CD4 belongs to the immunoglobulin superfamily of molecules which generally serve in recognition processes[2,3]. The sequence of CD4[4,5] indicates that it consists of a large (~370 residues) extracellular segment composed of four tandem immunoglobulin-like domains, a single transmembrane span, and a short (38 residues) C-terminal cytoplasmic tail. The first domain (D1) shares several features with immunoglobulin variable domains, but the sequence similarities between immunoglobulins and the other extracellular domains (D2, D3 and D4) are more remote.

In humans, CD4 can be subverted from its normal immunosupportive role to become the receptor for infection by the human immunodeficiency virus (HIV)[1,6,7]. Recombinant soluble CD4 proteins bind to the HIV envelope glycoprotein gp120, and can thus inhibit viral infection and virus-mediated cell fusion *in vitro* (refs 8, 9 and references therein). Domain D1 suffices for high-affinity binding to gp120 (ref. 8), and the analysis of substitution mutants further limits the sites of interaction to discrete regions in the domain[8,10–18].

Crystals of whole soluble CD4 (sCD4) molecules have been grown[19,20] but there has been limited success in achieving adequate diffraction order. The high solvent content and weak diffraction of several characterized polymorphs of human sCD4 are compatible with an extended, flexible molecule[19]. From the pattern of proteolytic cleavages that generate stable fragments (refs 21–23 and unpublished results), the main flexibility seems to be at the D2 to D3 junction. We have now crystallized a truncated derivative of CD4 that diffracts well, and here we report its atomic structure. This recombinant fragment[8] as secreted from Chinese hamster ovary (CHO) cells consists of residues 1–183 of human CD4 plus two missense residues, Asp-Thr; and it is unglycosylated. This molecule, which we refer to as D1D2, is as active as sCD4 in binding to gp120 (dissociation constant $K_d = 3$ nM) and retains all antibody epitopes mapped to these domains of CD4 (ref. 8 and unpublished results). Others have crystallized similar fragments from the N-terminal half of sCD4[24,25] and the structure of one is reported in the accompanying paper[25].

Here we describe the D1D2 structure in comparison with that of immunoglobulin domains, provide a geometrical definition for HIV recognition sites, and discuss implications of the structure for normal CD4 function and evolution of the immunoglobulin family. We find that the domains of CD4 are indeed immunoglobulin-like, although there are significant differences from the antibody analogues. The primary sites for HIV interaction are on loops that protrude from the variable-like D1 domain in analogy with immunoglobin complementarity-determining regions (CDRs). The D2 domain, which is intimately associated with D1, resembles constant domains but it is distinguished by a strand topology that is variable-like.

Structure determination

D1D2 protein secreted from CHO cells was purified to homogeneity by following procedures similar to those used for sCD4 (ref. 26). Crystals of this protein grown from polyethylene glycol (PEG) and stabilized at pH 8.2 (Table 1) belong to space group C2 and have unit cell dimensions of $a = 83.71$ Å, $b = 30.07$ Å, $c = 87.54$ Å, $\beta = 117.3°$. They contain one D1D2 molecule per asymmetric unit and 50% solvent. In searching for derivatives, we soaked crystals into stabilization medium doped with various heavy-atom compounds.

Diffraction data for the structure determination (Table 1) were measured at an area detector facility[27] where we collected CuKα data for the native protein and for candidate derivatives. Observable intensities could be recorded to 1.9 Å spacings from fresh native crystals, and even more strongly from the $K_2Pt(NO_2)_4$ derivative. We could also measure multiwavelength anomalous diffraction (MAD) data from this platinum derivative using characteristic gold emission lines which bracket the PtL_{III} absorption edge[28]. Phases were evaluated by both multiple isomorphous replacement (MIR) and MAD methods (Table 1). Electron density maps from MIR phasing (including anomalous data) at 2.7 Å resolution and from MAD phasing at 3.5 Å resolution both showed similar features. We then probabilistically combined MIR and MAD phase information to produce a combined map at 2.7 Å resolution (Fig. 1*a*) which showed distinct solvent channels and an apparent two-domain structure. A complete chain-tracing consistent with the amino-acid sequence for the first domain was possible, but the second domain remained difficult to interpret. By using a hand-drawn molecular envelope, we then performed solvent flattening and density truncation to improve the map further (Fig. 1*b*). This enabled us to trace the chain through the second domain. The C-terminal 12 residues (174–185) could not be seen.

A complete model was built into the solvent-flattened map as displayed by FRODO[29] program on a graphic workstation. Fragments from well-refined structures were fitted onto Cα guide points measured from a minimap and used as the starting point for building. Refinement made use of PROLSQ[30] and XPLOR[31] programs and included several manual rebuildings. Resolution of the analysis was gradually extended to 2.3 Å. After analysis of model 5 (Table 1) at $R = 0.208$, we discovered on comparing manuscripts that this model and one developed by Wang *et al.*[25] differed in the alignments of sequence in two strands (E of D1 and B of D2) and in conformations at residues 1, 103–108, 134–139 and 151–154. We then tested a model with residues 66–73 and 112–120 positioned in the former places of 64–71 and 114–122, respectively, and with the 134–139 region also rebuilt. Further refinement produced model 6 (Table 1) with an R value of 0.196 and stereochemical ideality typified by an r.m.s. deviation of 0.018 Å from ideal bond lengths. Study of this still incompletely refined model confirms the revised

TABLE 1 Statistics from the crystallographic analysis

(a) Diffraction data

Derivative	d_{min} (Å)	X-ray line /wavelength (Å)	No. of measurements	No. of reflections	Data coverage (%)	R_{merge}
Native	2.3	CuK_α/1.542	24,407	7,882	92	0.061
$K_2Pt(NO_2)_4$	2.5	CuK_α/1.542	40,130	6,757 (12,667)	99 (93)	0.057 (0.053)
	2.7	$AuL_{\beta1}$/1.084	33,914	5,314 (9,978)	97 (97)	0.069 (0.062)
	2.5	$AuL_{\beta2}$/1.068	43,005	6,639 (12,505)	97 (92)	0.072 (0.065)
	2.7	$AuL_{\alpha1}$/1.276	30,724	5,432 (10,210)	99 (99)	0.062 (0.054)
K_2PtCl_6	2.8	CuK_α/1.542	22,033	4,786 (8,751)	96 (88)	0.094 (0.086)

(b) MIR analysis

d_{min} (Å)	11.1	7.7	5.9	4.8	4.0	3.4	3.0	2.7	Total
$K_2Pt(NO_2)_4$									
r.m.s. f_H/E_{iso}	2.18	1.69	2.01	1.48	1.04	1.12	1.24	1.29	1.32
r.m.s. $\Delta F_{calc}/E_{ano}$	0.85	0.88	1.08	1.08	0.79	0.74	0.49	0.37	0.63
K_2PtCl_6									
r.m.s. f_H/E_{iso}	1.54	1.72	1.92	1.52	1.13	1.19	1.23	—	1.35
r.m.s. $\Delta F_{calc}/E_{ano}$	0.45	0.54	0.58	0.39	0.24	0.21	0.14	—	0.23
$\overline{m}$	0.81	0.71	0.75	0.71	0.58	0.57	0.51	0.37	0.53

(c) MAD analysis

		Observed ratio		
λ (Å)	1.542	1.084	1.068	1.276
1.542	0.043	0.066	0.057	0.069
1.084		0.036	0.029	0.030
1.068			0.038	0.036
1.276				0.037

(d) Solvent flattening

Shell/cycle	R	$\Delta\phi$
1/1	0.31	28.2
1/5	0.24	35.5
3/5	0.23	37.0
7/5	0.22	43.8
13/10	0.21	63.4

(e) Refinement

Step	Program	Data range (Å)	B-value refinement	Rebuilding	H_2O	R	Δd (Å)
0	PROLSQ	10.0–2.7	No	No	0	0.447	0.044
1	PROLSQ	5.0–2.7	No	No	0	0.271	0.025
2	PROLSQ	10.0–2.3	Yes	Yes	0	0.273	0.024
3	XPLOR	10.0–2.3	No	Yes	0	0.268	0.024
4	PROLSQ	10.0–2.3	Yes	Yes	0	0.250	0.023
5	PROLSQ	10.0–2.3	Yes	Yes	106	0.208	0.020
6	XPLOR/PROLSQ	10.0–2.3	Yes	Yes	72	0.196	0.018

Crystallization. D1D2 was purified in sequential steps of ion-exchange chromatography[28] involving S-Sepharose capture at pH 5.2, Q-Sepharose passage at pH 9.0 and a final elution from S-Sepharose at pH 6.0 in 0.6 M NaCl. Crystallization was carried out at 20 °C by equilibration against reservoirs of 20–25% PEG 3350 starting from hanging droplets composed of protein at −20 mg ml^{-1} in PBS at pH 7.0 and equal volumes from the reservoir. Collected crystals were transferred to a stabilization medium (30% PEG 3350, 80 mM Tris buffer, pH 8.2). Typical Crystals have dimensions of 0.5 × 0.3 × 0.2 mm. **Diffraction data** for MIR phasing and for refinement were measured with $CuK\alpha$ X rays generated by a Rigaku RU200 rotating anode and detected on two multiwire chambers on the Mark III system at UCSD. All data were collected at or near 20 °C. The inverse beam method was used to collect Bijvoet mates; that is, settings were changed from (ω, χ, ϕ) to $(\omega, -\chi, \phi + \pi)$ after rotations at each crystal orientation. Data for MAD phasing were collected on the Mark II area detector system. A graphite monochromator was used to select specific characteristic lines produced from a gold-plated anode installed on an Elliot GX-6 generator. The monochromator was tuned to 50% intensity at one side or the other of the unresolved $L_{\beta1}/L_{\beta2}$ doublet for selection of a primary monochromatic component. The UCSD area detector processing package was used for the reduction of all data. Values given inside of parentheses for numbers of reflections and R_{merge} are without merging of Friedel mates ($R_{merge} = \sum |I_{obs} - I_{avg}| / \sum I_{avg}$). **MIR analysis.** Heavy-atom positions were determined from Patterson syntheses. The refinement of heavy-atom parameters and the phase calculations were performed with a version of program PHARE modified by Z. Otwinowski (personal communication) for an improved error treatment. Relative occupancies for the $K_2Pt(NO_2)_4$ and K_2PtCl_6 derivatives were (1.00, 0.94) and (0.53, 0.61, 0.25, 0.22, 0.19), respectively. The first two heavy-atom sites of the K_2PtCl_6 derivative are in common with the sites of the $K_2Pt(NO_2)_4$ derivative. Parameters cited in the table are: f_H calculated heavy-atom structure-factor contributions; ΔF_{calc}, calculated Bijvoet difference; E_{iso}, r.m.s. isomorphous closure error; E_{ano}, r.m.s. anomalous closure error; and $\overline{m}$, mean figure-of-merit. **MAD analysis.** The MADSYS system of programs[54] was used to produce initial MAD phases. Phase combination between MIR and MAD was performed by the combination of ABCD coefficients[56]. But, in this case it was necessary to modify the original MAD refinement formulation[54] so as to produce phases for the native non-anomalously scattering component[56] of the platinyl CD4 complex. The previously used method for extracting ABCD coefficients[55] proved to be ineffective in this case; instead, these phasing coefficients were extracted directly from MAD phases and figures-of-merit[57]. MIR and MAD phases were combined at 20.0–3.5 Å; MIR phases were used alone for the 3.5–2.7 Å data. Observed anomalous diffraction difference ratios are given for the data between 20.0–3.5 Å. Bijvoet (diagonal elements of matrix) and dispersive (off-diagonal elements) difference ratios represent r.m.s. $(\Delta F_{\Delta\Delta})/$r.m.s. (F) and r.m.s. $(\Delta F_{\Delta\Delta})/$r.m.s. (F), respectively. **Solvent flattening** was performed following procedures used in the structure analysis of myohemerythrin[58]. After interpretation of the first domain, a molecular envelope was delineated which included 58% of the asymmetric unit of the crystal (as compared with 50% for the true solvent content) to ensure that all protein density was included. After subtraction of the mean solvent density and truncation of the lowest density values within the molecular boundary, the contents of the modified density function were Fourier-inverted. Phasing coefficients from the inversion were reduced by 0.5 before combination with ABCD values from MIR. A truncation level to eliminate the lowest 20% of molecular points (as compared with 0% and 10%) was chosen as giving the best definition for known features in domain D1. The phase refinement was carried out starting from the 20.0–4.5 Å shell, and including higher angle reflections successively until finally reaching 2.7 Å spacings. Five cycles of solvent flattening were carried out for each shell between shell 1 and shell 12, and 10 cycles were run for shell 13. R Crystallographic R value between observed structure factor magnitudes and calculated values from the solvent-flattened map. $\Delta\phi$ Phase differences between phases from MIR plus MAD and phases from the solvent flattened map. **Refinement.** Selected points along the course of refinement are listed. For all steps, the data having $F_{obs} > 4\sigma(F_{obs})$ were included. Step 0 refers to the unrefined starting model after adjustment of the scale factor only. From step 2 onward, the data between 10.0 and 5.0 Å were included which helped to achieve better connectivities in $2F_{obs} - F_{calc}$ maps. The final model at 2.3 Å resolution is based on the 5,689 reflections (64%) which met the 4σ criterion. This partially refined model includes the 1,420 atoms from residues 1–173 plus the 72 water sites. Thermal parameters were restrained as typified by an r.m.s. discrepancy of 1.17 Å in B values between bonded main-chain atoms. $R = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$; Δd, r.m.s. bond deviation from ideality (Å).

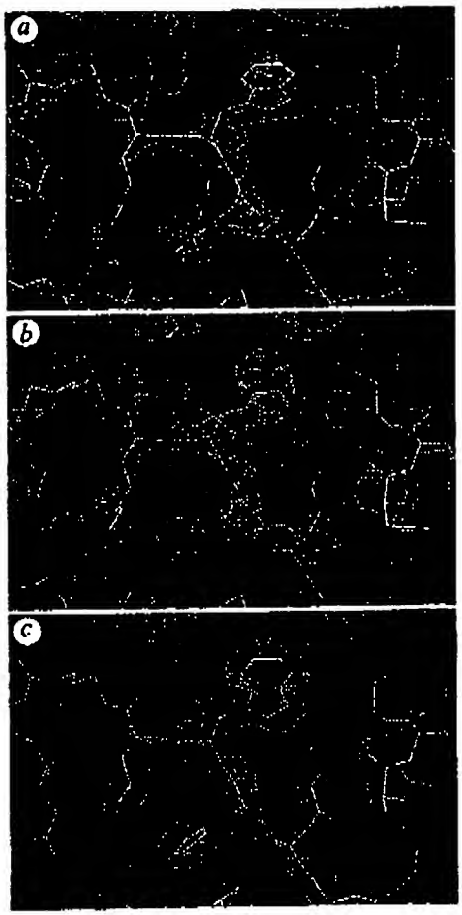


FIG. 1 Electron-density distributions used in the structural determination. The portion displayed in each panel includes the segment Phe 26-His 27-Trp 28-Lys 29 with the partially refined model 5 (yellow) superimposed on the density (blue). *a*, Experimental map at 2.7 Å resolution based on combined MIR plus MAD phase information; $\overline{m}=0.58$. *b*, Experimental map at 2.7 Å resolution after phase refinement by solvent flattening and density truncation; $\overline{m}=0.86$. *c*, Refined $2|F_{obs}|-|F_{calc}|$ map with model phases after refinement at 2.3 Å resolution; $R=0.208$.

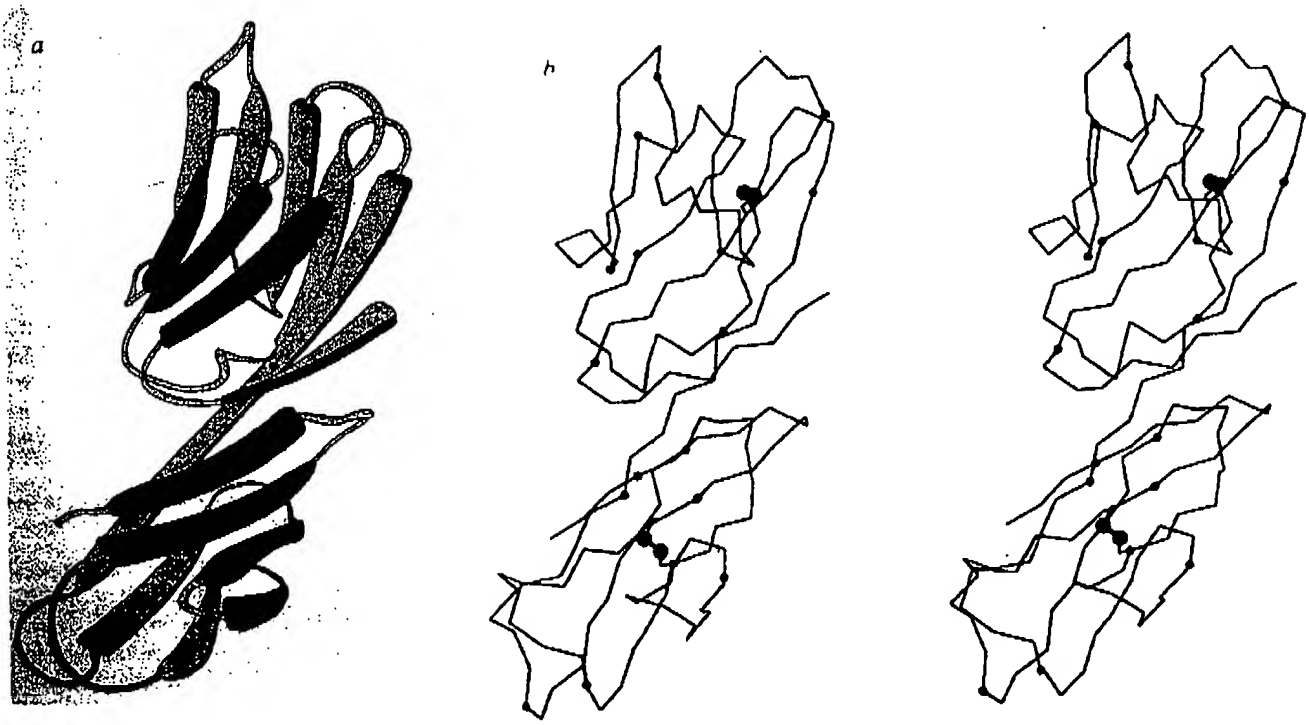
alignment in D1 and supports the changes made in D2, although several loops in D2 remain ill-defined. All conclusions drawn from model 5 remain unchanged, but the precise numerical quantities and Figs 2*b*, 4 and 5*a* were redone from model 6.

Overall structure

The D1D2 structure is of the all-β type. It consists of two intimately associated domains, each of which is a β sandwich folded with immunoglobulin-strand topology. The polypeptide folding is shown in Fig. 2. The first domain (D1) comprises residues 1–98 disposed in nine β strands, and the second domain (D2) contains residues 99–173 and has seven β strands. These domain boundaries are in striking correspondence with intron boundaries immediately after residues 100 and 177 in the gene structure[5]. The suggestion that J-like regions follow these introns is not borne out by the structure.

The last strand of D1 continues straight into the first strand of D2, running for a length of 49 Å from residues 88–103. This tandem association of domains leads to a rod-shaped molecule of dimensions roughly 25 × 25 × 60 Å. By comparing the solvent-accessible surface area of D1D2 with that of the separated

FIG. 2 Backbone structure of the D1D2 fragment of CD4. *a*, Schematic diagram (copyright Yarmolinski and Hendrickson, 1990). *b*, Stereodiagram of the α-carbon backbone. Positions of residue numbers divisible by ten are indicated by small spheres, sulphur atoms in disulphide bridges are indicated by larger spheres. The point of view is down the Z* axis after rotations about X(100°), Z'(−30°), Y'(−90°) and X*(10°) starting with the frame having X, Y, Z along a, b and c*.



domains, we found that 310 Å² from each domain is buried in the interface. This compares with 110 Å² for the VH-CH interface of Fab New[32]. Interdomain contacts are largely hydrophobic. Residues involved in interdomain contacts include Val 3, Leu 5, Ile 76, Leu 96, Val 97 and Phe 98 from D1, and Glu 119, Pro 121, Gln 163, Val 168 and Phe 170 from D2. The surface of D1D2 has many positively charged amino acids (the calculated pI of residues 1–183 is 10.0), and the electrostatic potential surface[33] computed from the model is mostly positive at neutral pH. There are, however, two prominent patches of negative potential: one is associated with Glu 85, Glu 87 and Asp 88 and the other is from Asp 105, Asp 153 and Asp 173.

TABLE 2 Residues implicated in HIV recognition by point mutation in D1 of CD4*

Amino-acid residue	Fractional solvent accessibility†	Effect on gp120 binding‡§				Effect on syncytium formation‡‖		Ref.
K29	0.30	A	■□□					16
G38	(0.31)	E	□□□	E	■■□	E	□□	10, 15
Q40¶	0.85	P	□□□	H	□□□	P	■□	10, 12
		H	□□	H	□□□			8, 15
		A	▨▨					16
G41	(0.93)	S	■□□					15
S42	1.00	Y	■□□	A	□□□			15, 16
F43	0.90	V	■■	L	■□□	V	■■	8, 15
		I	□□			L	□□	10
L44	0.38	S	■■□	A	■■□			15, 16
T45	0.69	P	■■■	G	■□□	P	■■	10, 15
		I	□□	A	□□□	I	□□	8, 16
G47	(0.43)	E	□□	S	■■□	E	■□	10, 15
		R	■■■			R	■■	10
S49	0.58	T	□□□	Y	■■□			15
		A	■□□					16
N52	0.71	K	□□□	A	■□□			15, 16
A55	0.00	F	■■			F	■■	8
R58	0.73	G	■□□	A	□□□			15, 16
R59	0.82	G	■□□	A	■■□			15, 16
Q64	0.84	E	□□□	A	■□□			15, 16
		K	□□					12
F67	0.00	L	■■□					15
E77	0.97	A	■□□					16
T81	0.56	A	■■■					16
E85	0.59	A	■■□					16
E87	0.53	K	□□			K	■□□□	10, 17
		A	□□□			G	■■■■	16, 17
D88	0.71	A	□□□			N	■□□□	16, 17
Q89	0.89	L	□□			L	■■■□	10, 17
		A	□□□			K	■■□□	16, 17

* Within the 98 amino-acid residues of the D1 domain, there is no mutational information on residues 3–6, 12, 14, 16, 18, 26, 69–71, 76, 83, 84 and 97. However, the binding of gp120 to Rhesus CD4[10] and to the human-rat chimaeric protein eliminates 12, 14, 18 and 76 from this list. Mutations which caused global alteration of structure, as evidenced by reduced binding of several anti-CD4 monoclonal antibodies, were not considered[15,16]. Such mutations were the only data for 13 amino acids and thus led to exclusion of residues 7, 13, 28, 36, 37, 54, 57, 62, 78, 79, 82, 93 and 95. Thus, in total, 25 amino acids are excluded from consideration.

† Fractional solvent accessibility was calculated as the ratio of solvent-accessible surface area for atoms of an amino-acid residue X in the protein to that area obtained after reducing the structure to a Gly-X-Gly tripeptide[39]. Values cited are for side-chain atoms except in the case of glycine residues where main-chain values are given.

‡ Mutants are identified by single-letter code. Mutational effects are symbolized by blackening in steps proportional to the quantitational sensitivity of the assays as reported. Only single amino-acid substitutions are included. The results from double substitutions largely confirm these data. Certain double mutations are of particular note: A, G41S/F43C severely disrupts binding[16], emphasizing the significance of F43; B, L51M/A55S does not affect binding[16], indicating that the severe effect by the A55F substitution is conformational in nature; C, E87G/K90G[13] and D88N/Q89R[8] do not affect binding, further indicating that this region is not involved in initial binding to g120.

§ Relative effects of the indicated mutations on virus[10] or gp120[15] binding to cell surface CD4 or gp120 binding to soluble CD4 proteins[8,12,16]. The results shown are an interpretation of the primary data and comparisons among the results from the different laboratories are approximate.

‖ Relative ability of cells bearing CD4 mutant proteins to form syncytium with cells expressing viral envelope proteins[10,17] or of soluble CD4 proteins to inhibit syncytium[8].

¶ The Q40A mutation increased binding affinity to gp120 twofold.

The D1D2 crystal lattice contains diad axes that could accommodate a dimeric molecule. However, contacts across screw axes dominate the lattice interactions and the only diad-mediated interface (between D1 domains) seems unlikely to persist in solution. The lattice interactions of D1 are more extensive than those of D2 which is reflected in molecular mobilities of the domains: the average B value is 19.9 Å² for D1 compared with 37.8 Å² for D2. This accounts for our greater difficulty in tracing the D2 chain.

CD4, like many members of the immunoglobulin superfamily, is a single-chain cell-surface receptor composed of tandemly repeated domains. Some of these domains have been assigned as V-like and others as C-like or in the 'C2 set'[34]. The D1D2 domain of CD4 contains both types and, apart from the similarity of CD5 with the PapD bacterial chaperone protein[35], this is the first three-dimensional structure for such members of the family.

Variable-like domain D1

As anticipated by sequence comparisons, D1 is similar to the variable (V) domains of antibodies. A schematic drawing of the folding pattern with standard immunoglobulin designations for the secondary structural elements is shown in Fig. 3. Strands B, D and E make up one β sheet, and strands A, C, C′, C″, F and G are in the other sheet of the β sandwich. In immunoglobulin V domains, the first half of strand A is hydrogen-bonded to strand B of the outer sheet but the second half switches to the inner sheet. Strand A of D1 is foreshortened and occupies only the 'inner-sheet' position, hydrogen-bonded in parallel with strand G. D1 preserves the inter-sheet disulphide bridge and several other elements of the hydrophobic core characteristic of the immunoglobulin framework. Starting from these alignments we have superimposed D1 onto representative immunoglobulin V domains from Bence-Jones protein Rei (VLκ)[36], and Fab New (VLλ and VH)[32]. The structurally aligned sequences are shown in Fig. 4. D1 superimposes remarkably well onto the β-sheet framework of all V domains, with a best match to Rei bringing 72 Cα atoms to an r.m.s. discrepancy of 1.22 Å from exact superposition.

In striking contrast with the conserved core of β strands, several loops between strands in CD4 are quite different from those in immunoglobulin V domains. In particular, loop CC′ is shortened by four residues and loop FG (immunoglobulin CDR3) is shortened by four to six residues from canonical immunoglobulin lengths. These loops mediate VH–VL dimerization in immunoglobulins, an interaction that does not occur with CD4. One addition, important for gp 120 binding, is the lengthening of loop C′C″ (immunoglobulin CDR2) by three residues in CD4 compared with κ light chains such as Rei. In this regard, CD4 more closely resembles VH domains even though overall it superposes best with Rei. This CDR2-like loop interacts with Trp 62, not found in antibodies, and juts out at the tip. Few of the loop changes in D1 were detected in previous alignments with Rei, and consequently structural predictions based on Rei[37] have failed to capture the essence of CD4. The salience of the CDR2 analogue of CD4 and the diminutive nature of the CC′ loop and the CDR3 analogue are evident in Figs 2 and 3.

Distinctive domain D2

Domain D2 of CD4 has a tertiary folding that can readily be recognized as resembling immunoglobulin constant domains. Again following standard immunoglobulin nomenclature (Fig. 3), one β sheet of the sandwich-like structure contains strands A, B and E, and the other sheet has strands C, C′, F and G. Despite the general similarity of D2 to constant domains, details of the folding are considerably divergent. First, the size of D2 is small (75 residues) compared with that of immunoglobulin constant domains (~100 residues). This manifests itself in shorter strand lengths. A second variation is that strand C′.

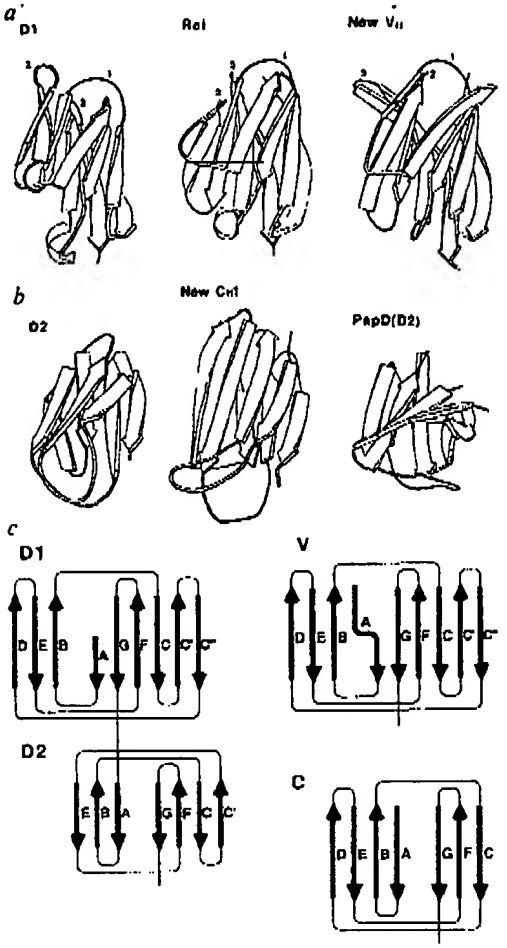


FIG. 3 Schematic comparisons between domains of CD4 and other immunoglobulin-folded domains. *a*, Ribbon diagrams comparing D1 with a variable λ light-chain domain and a variable heavy-chain domain. The point of view is as in Fig. 2; the CDR loops of immunoglobulins and their analogues in D1 are indicated by numerals. *b*, Ribbon-diagram comparisons of D2 with a constant domain and with domain D2 of the bacterial chaperone protein PapD. *c*, Topology diagrams and strand nomenclature for the β sheets in CD4 and in variable and constant domains of immunoglobulins.

which corresponds sequentially to the D strand of normal constant domains, joins the β sheet consisting of strands C, F and G, rather than being hydrogen-bonded to strand E in the other sheet. In this respect, D2 is V-like. A final striking difference from immunoglobulins is that D2 has its disulphide bond between strands in the same sheet, that is strands C and F (Fig. 2), rather than between sheets as is usual. This feature, which is unusual but not unprecedented[38], had been predicted[34]; however, strand C' was not anticipated. The hydrophobic core of immunoglobulin domains is preserved in D2 despite the unusual disulphide connection (Fig. 4). Leu 116 on strand B occupies the position of a normal cysteine partner for Cys 159 on strand F. The inward orientation of the intra-sheet disulphide bridge makes it an important member of the hydrophobic core even though it does not join the sheets.

D2 can be structurally aligned with the CH1 domain of Fab New (Fig. 4), but the resulting spatial superposition is not as close as for D1 comparisons with V domains. Only six of the seven strands superimpose at a 2.5 Å stringency for matches, and this gives 33 matches with an r.m.s. discrepancy of 1.59 Å. D2 superimposes somewhat less well with the Rei variable domain (30 matches within 2.5 Å for 1.68 Å r.m.s. discrepancy). In fact, D2 is actually more similar to the domains of PapD[35] than to immunoglobulin domains. As in the D2 topology, PapD domains also exhibit sheet switching (partially in D1(PapD), fully in D2(PapD)), and 39 Cα atoms of D2 (CD4) superimpose on D1(PapD) with an r.m.s. discrepancy of 1.55 Å. Drawings of the aligned structures are shown in Fig. 3.

Binding site for gp120

The initial molecular event in HIV infection is the binding of gp120 on the viral envelope to CD4 on the cell surface. The high affinity of this interaction, at least for several laboratory strains of the HIV-1 virus[39], has permitted a detailed mapping of the binding site. More than 200 mutant CD4 recombinants have been constructed and tested for gp120 affinity (refs 8, 10-18, J. Arthos *et al.*, manuscript submitted, and unpublished results). Some of these mutant proteins have also been used to map the CD4 epitopes of a battery of monoclonal antibodies. In Table 2 we distil the mutation results to those point mutants in D1 which impair gp120 binding without causing extensive disruptions of structure as judged by antibody binding. Analyses have been reported for mutations at 80 of the 98 residues in D1, but only 19 positions among them have impact on gp120 binding without apparent global conformational change. Thirteen of these are in the span from 38 to 59. Completely buried

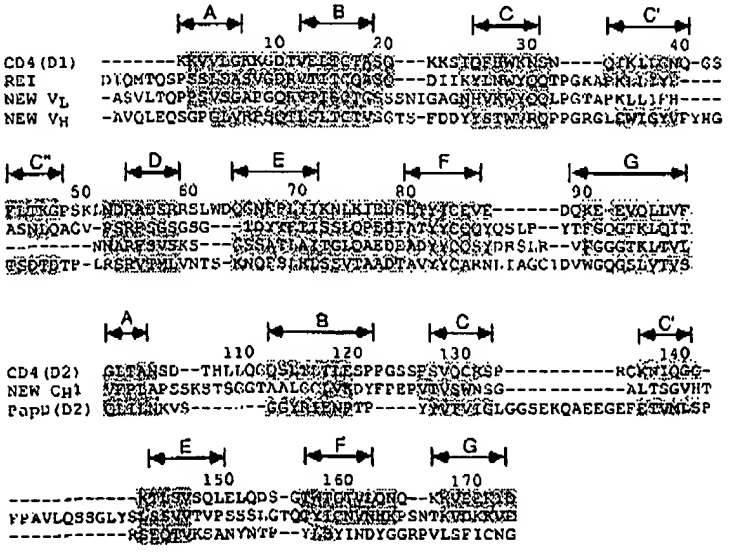


FIG. 4 Structural alignment of the amino-acid sequences of CD4 domains with other immunoglobulin-related domains. Shaded residues have Cα positions within 2.5 Å of corresponding CD4 positions after optimal superposition of all shaded residues for a given pair of domains. (Exceptions up to 2.60 Å were allowed for residues in the middle of strands.) Each superposition relates a certain number of positions, *N*, within the specified 2.5 Å limit, and these match at a certain r.m.s. discrepancy, Δ. For the match of D1 with Rei ($V_L\kappa$), $N=72$ and $\Delta=1.22$ Å; for D1 versus New ($V_L\lambda$), $N=66$ and $\Delta=1.12$ Å; and for D1 versus New (V_H), $N=68$ and $\Delta=1.13$ Å. For the match of D2 with New (C_H1), $N=33$ and $\Delta=1.59$ Å; for D2 versus PapD(D2), $N=32$ and $\Delta=1.26$ Å; for D2 versus PapD(D1), $N=39$ and $\Delta=1.55$ Å; and for D2 versus Rei ($V_L\kappa$), $N=30$ and $\Delta=1.68$ Å. The alignments of D2 versus PapD(D1) and D2 versus Rei ($V_L\kappa$) are not shown.

residues can be discounted as direct participants in binding because their exposure would require major conformational change, which is without precedent for immunoglobulin domains. We have calculated the degree of exposure for each residue in D1 (Table 2), and find that three of the gp120-sensitive residues, Gly 38, Ala 55 and Phe 67, are inaccessible. It seems unlikely that either Lys 29 or Gln 64 is critical for binding because there is no effect for mutations at the more exposed neighbours of Lys 29 or for single or multiple mutations at 64, except alanine. Thus only the alanine replacements at 77, 81 and 85 fall outside of the 41–59 span. The distribution of binding mutants is shown in Fig. 5*a*, *b*.

The importance of the 41–59 region in HIV binding is corroborated by a chimaeric CD4 in which the human sequence from 36 to 62 has been inserted into rat CD4 (A. F. Williams, personal communication). This imparts full gp120 binding activity on the normally unreactive rat CD4 molecule. It also shows that 33 additional side chains which differ in rat and human CD4, including Gln 64, are not absolutely essential for the binding[4]. But, Glu 77, Thr 81 and Glu 85 are among the residues conserved between rodents and primates. Thus, whereas the 41–59 region is positively identified as being involved in gp120 binding, participation of the 77–85 region cannot be excluded.

Antibody binding studies also help to define the gp120 binding site (refs 10, 40, and our unpublished results), pointing especially to the CDR2-like region. Two major groups of D1-binding monoclonal antibodies have been described and their epitopes are illustrated in Fig. 5*c*. The epitope of one cross-blocking group, typified by Leu3a, maps to the CDR2-like loop and usually includes portions of the CDR1-like loop. These antibodies also block gp120 binding. The epitope of the second group, typified by monoclonal L71, includes the CDR3-like loop and sometimes a portion of the strand G. These antibodies inhibit viral infection and cell fusion, but do not efficiently block gp120 binding; some even form ternary complexes with gp120 and CD4 (A. Trunch *et al.*, manuscript submitted).

The CDR2-like region that is strongly implicated in gp120 binding is also very prominent in the CD4 structure (Figs 3 and 5). The C′–C″ hairpin loop at Gln 40–Phe 43 is almost completely exposed. Most strikingly, the hydrophobic side chain of Phe 43 juts out into the solvent as shown in Fig. 5*e*, *f*. The involvement of a phenyl group in binding is also suggested by peptide inhibitor studies[41,42]. If the 77–85 span were also to be directly involved in gp120 binding, this would present a puzzle because this site is on a face nearly opposite from the CDR2-like region (the vector from the molecular centre of D1 through Cα43 makes an angle of 152° with the similar vector through Cα81), and no intervening residues have been implicated in gp120 recognition.

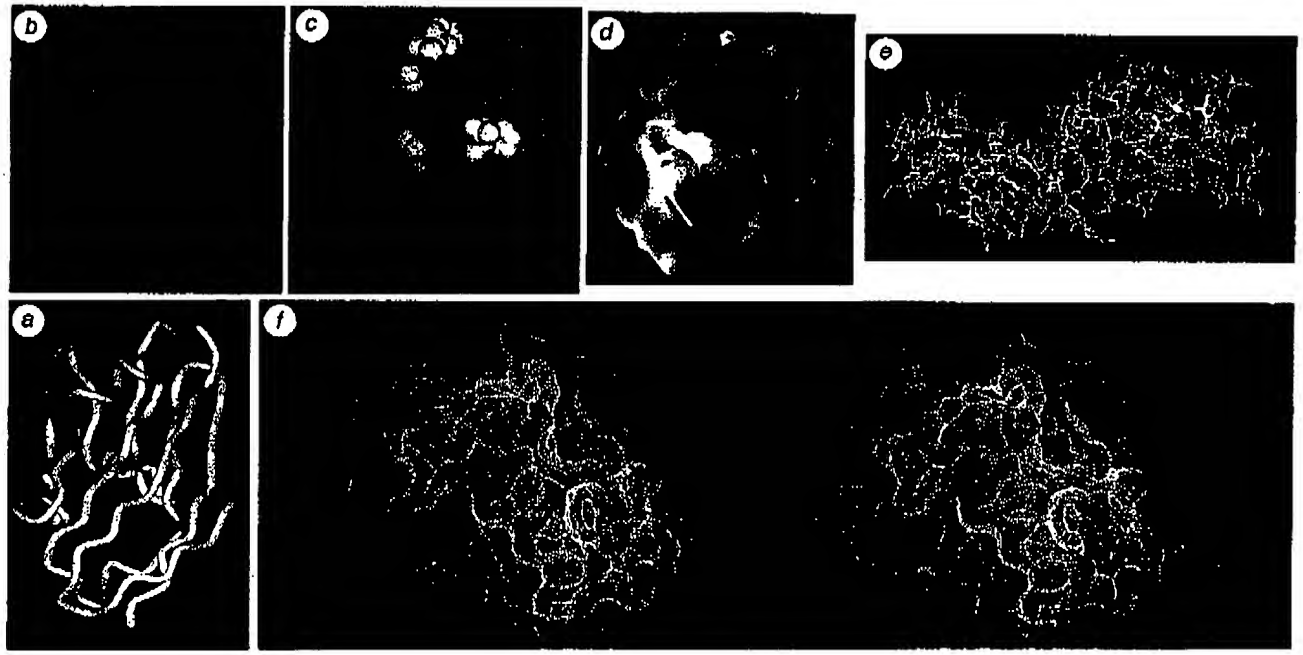


FIG. 5 Images of CD4 relating to sites of interaction with HIV. *a*, Cα backbone of D1 with sites implicated by mutational and structural analysis to affect high-affinity gp120 binding (see Table 2). Non-buried residues that show reduction in binding without global disruption of structure (as judged by antibody binding) are coloured red-orange (29, 41, 42, 43, 44, 45, 47, 49, 52, 58, 59, 77, 81 and 85) on a backbone drawn by WORM (L. Andrews). The orientation is as in Figs 2 and 3. *b*, Van der Waals' surface of D1 with side chains of gp120-sensitive residues coloured. Those showing marked reductions are in pink (43, 44, 85, are visible) and those showing moderate effects are in purple (59, 42, are visible). The point of view is from above *a*, looking down at the tips of CDR-like loops. *c*, Van der Waals' surface of D1 with side chains of antibody epitopes coloured. Residues identified with epitopes of the Leu3a family and shown in purple (24, 25, 27, 42, 43, visible) and those identified with the L71 family are shown in yellow (88, 89, visible). The view is roughly as in *b*. *d*, Electrostatic potential surface computed with DELPHI[33] at neutral pH, and displayed with AAK (A. Nicholls and B. Honig) at the levels of the solvent-accessible surface with a probe radius of 1.4 Å. Blue represents positive potential, red negative, and white neutral. The negative patch is associated with residues 85, 87 and 88. The view is approximately as in *b*. *e*, An all-atom representation of D1D2. Residues in the gp120-binding region from residue 41 to 59 are drawn in red. The direction of view is roughly as in Fig. 2, but the molecule has been rotated by ~90° about this view axis. This view illustrates the exposed nature of Phe 43. The actual conformation of this phenyl side chain in the crystal is at least partly determined by lattice interactions, but its highly exposed nature would also be expected to persist in other conformations accessible in solution. *f*, Stereoview of the molecular surface in the major gp120-binding region of D1. Atoms are drawn as in *e*, and are enveloped by the surface in contact with a probe sphere of 1.4 Å radius as displayed by QUANTA (Polygen). The view is taken after rotation from *e* by ~90° about the horizontal axis.

Fusion determinants

After the initial binding event, entry of virus into the cell occurs through fusion of cellular and viral envelope membranes. In addition, HIV-infected cells fuse with uninfected CD4$^+$ cells to form syncytia, a process mediated by interactions between HIV envelope protein expressed on the infected cell surface and CD4 on the uninfected cell. Syncytium formation is thought to mimic viral entry and could be an important mode of viral spreading. Chimpanzee CD4, which binds to HIV but does not support syncytium formation, has Gly replacing Glu at residue 87. This replacement in human CD4 abolishes syncytia formation while preserving gp120 binding[17] and, curiously, normal viral infection kinetics. Residues 88 and 89 are also implicated in this process by mutation (Table 2). These residues lie at the β-turn tip of the CDR3-like loop in D1 and are part of a patch of negative potential on the otherwise mostly positive D1D2 surface (Fig. 5*d*). The CDR3-like loop is spatially separated from the CDR2-like loop implicated in high-affinity binding to gp120 (Cα87 is 17 Å from Cα43). In accord, monoclonal antibodies of the L71 family block syncytium formation although they permit gp120 binding to CD4. So, it seems that the CDR3-like loop is necessary for secondary interactions between the viral envelope and CD4 before fusion. This could relate to CD4-induced release of gp120 from virus and infected cells[43-45] and exposure of a fusogenic domain on the membrane envelope protein gp41.

Immune response interactions

The observation that CD4$^+$ lymphocytes react only with class II-bearing target cells suggests that CD4 associates directly with class II MHC, and it is thought that CD4 may also have a role in signal transduction. There is evidence from cellular assays to support the association of CD4 with class II MHC molecules[46], the T-cell receptor[47], and the p56^{lck} kinase[48], as well as other CD4 molecules[49]. Except for the interaction of CD4 with p56^{lck}, however, these associations are necessarily weak and thus difficult to measure. Mutagenesis *in vitro* has been used to identify the sites on CD4 that interact with class II MHC molecules. On the basis of assays of either cell adhesion[18,50] or interleukin-2 production[18], large separated expanses of the CD4 molecule, involving regions on D1, D2 and D3, have been implicated in the CD4–MHC interaction. These findings are not easy to reconcile with the structure of CD4. Some of the disruptive mutations occur in contact or bridging regions between the domains. Perhaps some of the other disruptions reflect complications from the diverse components of the cellular system rather than direct binding. For example, self-associations of CD4[19] might be involved in signal transduction and possibly in class II binding.

Evolutionary implications

There is little doubt that CD4 and immunoglobulins have evolved from a common ancestor. Intron and domain boundaries coincide and, although the sequences are highly divergent, superimposable strand topologies and a common hydrophobic core are preserved. It seems likely, however, that evolution to the antibody family, with its vast repertoire of dimeric receptor units, was a later event, and that CD4 is more prototypical of immunoglobulin superfamily receptors which are often monomeric and nonpolymorphic. The absence from CD4 of J-like regions, which impart diversity in antibodies, is consistent with this. The progenitor of the diverse immunoglobulin superfamily of the present may have its vestige in the gene structure of D1, which is split by an intron (at position 47) as are genes for other immunoglobulin-like domains[51,52]. A quasi-diad axis, perpendicular to the sheets and passing midway between strands B and E on one sheet and between strands C and F on the other, can be used to superimpose successive strands in the first exon on those in the second exon. An immunoglobulin progenitor produced by a gene duplication event[53] might have been V-like, but a D2-like progenitor that would evolve to V and C domains is also a possibility.

Whatever the course of evolution, it is remarkable that molecules that function in recognition are often members of the immunoglobulin gene family. Clearly the immunoglobulin fold provides a facile framework from which a variety of loops with distinctive characteristics can be elaborated. These loops can have distinct functions, as in binding to gp120 at the CDR2-like loop and affecting fusion from the CDR3-like loop, and such modularity might have evolutionary advantage. Similarly, the quasi-symmetric nature of the immunoglobulin motif, with N and C termini at opposite ends, facilitates the concatenation of tandem, flexibly linked modules that can have distinct roles. In the case of CD4, one domain might be involved in class II binding whereas another domain might effect self-associations of the kind found in crystalline polymorphs[19], and these could be essential for signal transduction. Indeed, in an alignment of D3 with D1, dimerization loops of immunoglobulin V domains would not be foreshortened in D3 as they are in D1. Such modules could evolve separately and be recombined by exon shuffling to integrate complex recognition and effector functions. □

Received 12 October; accepted 29 October 1990.

1. Robey, E. & Axel, R. *Cell* **60**, 697-700 (1990).
2. Maddon, P. J. *et al. Cell* **42**, 93-104 (1985).
3. Williams, A. F. & Barclay, A. N. *A. Rev. Immun.* **6**, 381-405 (1988).
4. Clark, S. J., Jefferies, W. A., Barclay, A. N., Gagnon, J. & Williams, A. F. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1649-1653 (1987).
5. Maddon, P. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 9155-9159 (1987).
6. Dalgleish, A. G. *et al. Nature* **312**, 763-767 (1984).
7. Klatzmann, D. *et al. Nature* **312**, 767-768 (1984).
8. Arthos, J. *et al. Cell* **57**, 469-481 (1989).
9. Capon, D. J. *et al. Nature* **337**, 525-531 (1989).
10. Peterson, A. & Seed, B. *Cell* **54**, 65-72 (1988).
11. Landau, N. R., Warton, M. & Littman, D. R. *Nature* **334**, 159-162 (1988).
12. Clayton, L. K. *et al. Nature* **335**, 363-366 (1988).
13. Mizukami, T., Fuerst, T. R., Berger, E. A. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9273-9277 (1988).
14. Chao, B. H. *et al. J. biol. Chem.* **264**, 5182-5188 (1989).
15. Brodsky, M. H., Warton, M., Myers, R. M. & Littman, D. R. *J. Immun.* **144**, 3078-3086 (1990).
16. Ashkenazi, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 7150-7154 (1990).
17. Camerini, D. & Seed, B. *Cell* **60**, 747-754 (1990).
18. Lamarre, D. *et al. Science* **245**, 743-746 (1989).
19. Kwong, P. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 6423-6427 (1990).
20. Davis, S. J. *et al. J. molec. Biol.* **213**, 7-10 (1990).
21. Richardson, N. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 6102-6106 (1988).
22. Ibagbu, C. C. *et al. J. Immun.* **142**, 2250-2256 (1989).
23. Healey, D. *et al. J. exp. Med.* **172**, 1233-1242 (1990).
24. Chamow, S. M. *Biochemistry* **29**, 9885-9891 (1990).
25. Wang, J. *et al. Nature* **348**, 411-418 (1990).
26. Carr, S. *et al. J. biol. Chem.* **264**, 21285-21295 (1989).
27. Xuong, Ng.h., Sullivan, D., Nielsen, C. & Hamlin, R. *Acta crystallogr.* **B41**, 267-269 (1985).
28. Ashford, V., Dai, X., Nielsen, C., Sullivan, D. & Xuong, Ng.h. *Acta crystallogr.* **A46**, C-15 (1990).
29. Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
30. Hendrickson, W. A. & Konnert, J. H. in *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Venkatesan, K.) 13.01-13.23 (Indian Institute of Science, Bangalore, 1980).
31. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458-460 (1987).
32. Saul, F. A. *et al. J. biol. Chem.* **253**, 585-597 (1978).
33. Gibson, M., Sharp, K. & Honig, B. *J. comp. Chem.* **9**, 327-335 (1987).
34. Williams, A. F., Davis, S. J., He, Q. & Barclay, A. N. *Cold Spring Harb. Symp. quant. Biol.* **54**, 637-647 (1989).
35. Holmgren, A. & Brändén, C.-I. *Nature* **342**, 248-251 (1989).
36. Epp, O. *et al. Eur. J. Biochem.* **45**, 513-524 (1974).
37. Bates, P. A., McGregor, M. J., Islam, S. A., Sattentau, Q. & Sternberg, M. J. E. *Protein Eng.* **3**, 13-21 (1989).
38. Wilson, I. A., Skehel, J. J. & Wiley, D. L. *Nature* **289**, 366-373 (1981).
39. Daar, E. S., Xi, L. L., Moudgil, T. & Ho, D. D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6574-6578 (1990).
40. Sattentau, Q. J. *et al. J. exp. Med.* **170**, 1319-1334 (1989).
41. Lifson, J. D. *et al. Science* **241**, 712-716 (1988).
42. Finberg, R. W. *et al. Science* **249**, 287-291 (1990).
43. Kirsh, R. *et al. AIDS Res. hum. Retroviruses* **6**, 1209-1212 (1990).
44. Moore, J. P., McKeating, J. A., Weiss, R. A. & Sattentau, Q. J. *Science* (in the press).
45. Hart, T. K. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
46. Doyle, C. & Strominger, J. L. *Nature* **330**, 256-259 (1987).
47. Chuck, R. S., Cantor, C. R. & Tse, D. B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6021-6025 (1990).
48. Turner, J. M. *et al. Cell* **60**, 755-765 (1990).
49. Veillette, A., Bookman, M. A., Horak, E. M., Samelson, L. E. & Bolen, J. B. *Nature* **338**, 257-260 (1989).
50. Clayton, L. K., Sieh, M., Pious, D. A. & Reinherz, E. L. *Nature* **339**, 548-551 (1989).
51. Owens, G. C., Edelman, G. M. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 294-298 (1987).
52. Lemke, G., Lamar, E. & Patterson, J. *Neuron* **1**, 73-83 (1988).

53. Bourgois, A. *Immunochemistry* **12**, 873-876 (1975).
54. Hendrickson, W. A. *Trans. Am. crystallogr. Assoc.* **21**, 11-21 (1985).
55. Parter, A., Smith, J. L. & Hendrickson, W. A. *Acta crystallogr.* **A46**, 537-540 (1990).
56. Karle, J. *Int. J. quant. Chem. Symp.* **7**, 357-367 (1980).
57. Hendrickson, W. A. *Acta crystallogr.* **B27**, 1472-1473 (1971).
58. Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2160-2164 (1975).
59. Sheriff, S., Hendrickson, W. A., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1104-1107 (1985).

ACKNOWLEDGEMENTS. We thank M. Skinner, D. Sullivan, V. Ashford and C. Nellsen for help in data collection; colleagues at SmithKline Beecham for help in protein preparation and purification; C. Ogata for helpful discussions; Z. Otwinowski for his phasing program; C. Brändén and M. Amzel for supplying PAPD and Feb New coordinates, respectively; A. Williams, Q. Sattentau, P. Bjorkman and D. Capon for communicating results before publication, and R. Yarmolinski for her drawing. This study was supported in part by the National Institutes of Health (W.A.H.) and by the Markey Foundation (N.-h.X.); the area detector facility at UCSD is an NIH Research Resource (N-LX). Atomic coordinates have been deposited in the Protein Data Bank with entry name 1CD4 for release one year after publication of this paper.

LETTERS TO NATURE

Entropy production as the selection rule between different growth morphologies

Adrian Hill

Physiological Laboratory, Downing Street, Cambridge CB2 3EG, UK

CRYSTALLIZATION of a solid phase from a melt or solution is a special case of pattern formation in which the dissipation of energy across the free energy gradient between the two phases can give rise to various growth morphologies in the steady state[1]. Experimental studies of crystallization from undercooled solutions[2], electrolytic deposition[3,4] and the formation of fluid patterns in a Hele-Shaw cell[5] have revealed faceted, dendritic, 'dense-branching' and fractal morphologies[6]. For a system with fixed anisotropy and interfacial tension, changes in the driving force for the transition (such as the degree of undercooling) can cause changes in growth morphology which are usually accompanied by changes in growth rate. The selection rule that determines these morphologies remains unclear, although a recent suggestion[5,6] is that it is based on the growth velocity. Here I propose that selection is governed by the rate of entropy production per unit area of the different growth patterns. This principle allows accurate prediction of the morphology transition observed for the crystallization of NH_4Cl (ref. 2). I suggest that it may reflect a more general thermodynamic principle underlying a wide range of natural processes.

Consider the formation of a generalized 'crystal' in which the driving force X (ignoring surface terms) is proportional to either supersaturation, pressure or other fields. The form of the crystal involves many surface orientations which have a different surface free energies. Denoting this surface free energy f by

$$f = u - Ts \tag{1}$$

where u and s are the surface energy and entropy of the crystal, the velocity V representing the rate of crystallization is expressed as a linear function

$$V = L(X - \theta) \tag{2}$$

where X is the difference between the free energies of the dissolved solute and the bulk solid and θ is the free energy required to increase the area, so that $X - \theta$ is the total driving force, and L is the rate coefficient reflecting the growth rate of a particular morphology and its rate of entropy production. The dissipation is given by

$$\phi = dS/dt = V(X - \theta) = L(X - \theta)^2 \tag{3}$$

where θ is a linear function of f and S is the entropy created in the crystallization. A model of the process of growth that relates velocity to driving force is a model of the rate coefficient and a link between the kinetics and the entropy production. Ben-Jacob and Garik[6] suggest that there must be a conjugate variable to the average rate of crystal growth. This conjugate variable is the force $(X - \theta)$. Equations (2) and (3) assume linear phenomenological relations but these are merely convenient and not essential: they almost certainly do not hold far from equilibrium or when chemical reactions occur.

The dissipation functions of two morphologies formed by adjacent steady states will in general have crossover points (Fig. 1). It follows directly from equation (3) that the crossover point for adjacent morphologies occurs at the driving force

$$X_c = \frac{(L_1\theta_1 - L_2\theta_2) + (\theta_1 - \theta_2)\sqrt{(L_1L_2)}}{(L_1 - L_2)} \tag{4}$$

There are entropic crossovers because equation (3) is a parabola in the $\phi - X$ plane: the section between $X = 0$ and $X = \theta$ represents a disaggregation of the crystal structure. This does not normally occur in experiments unless crystals are initially present, in which case one morphology disappears as the other is formed. As ϕ is always positive, I consider for the moment only the parts of the entropy curves with positive gradients because these represent crystallization, that is, the forward reaction driven by X. Reference to Fig. 1 shows that if a crossover point is to occur in this domain when $\theta_1 < \theta_2$ then $L_1 < L_2$. At higher driving forces new crossovers occur if there are steady states that can exist by creating other morphologies with higher values of L. When the driving force is reduced from a high value, the principle of selection for the steady state with highest value of ϕ predicts that at the crossover points (c_1 and c_2 in Fig. 1) transitions occur to different steady states producing different crystal morphologies. In the linear regime these transitions cannot occur without discontinuities in the rate of crystallization because, as follows from equations (2) and (4), the only point free from such discontinuity is the trivial one

$$(\theta_1 - \theta_2)\sqrt{(L_1L_2)} = 0 \tag{5}$$

in which the rates are zero. Small differences in entropy production rates at a crossover, however, could lead to small changes in slope together with small discontinuities which may be masked by variation in the data. Discontinuities are indeed a striking feature of morphology changes in various experimental

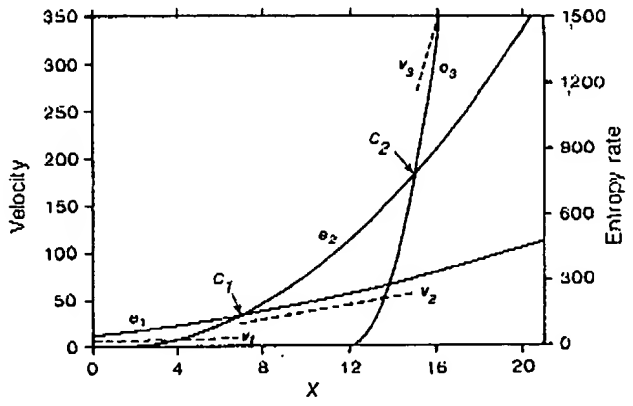


FIG. 1. The crossing points c_1 and c_2 of the entropic curves e_1-e_3 (solid lines) of three morphologies. The associated velocities v_1-v_3 (dashed lines) are also shown. The constants for the three entropic curves are $L_1 = 0.5$, $\theta_1 = -10$; $L_2 = 4$, $\theta_2 = 1$; $L_3 = 87$, $\theta_3 = 12$.